



CP 2623  
PATENT

Case Docket No. ABOHM1.001CP1  
Date: October 31, 2001

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants : Frommel et al.  
Appl. No. : 09/772,538  
Filed : January 29, 2001  
For : DETERMINATION OF  
LIGANDS FOR PROTEINS  
Examiner : Unknown  
Group Art Unit : 2621

I hereby certify that this correspondence and all  
marked attachments are being deposited with the  
United States Postal Service as first class mail in  
an envelope addressed to: Assistant Commissioner  
for Patents, Washington, D.C. 20231, on

November 2, 2001

(Date)

Kirk E. Pastorian, Reg. No. 48,756

RECEIVED

JAN 14 2002

Technology Center 2600

TRANSMITTAL LETTER

ASSISTANT COMMISSIONER FOR PATENTS


WASHINGTON, D.C. 20231

ATTENTION: APPLICATION BRANCH

Dear Sir:

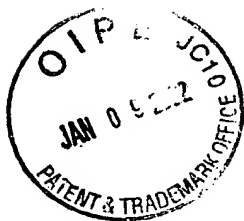
Enclosed for filing in the above-identified application are:

- (X) A certified copy of German Application No. DE 198 31 758.1-52.
- (X) The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment, to Account No. 11-1410.
- (X) Return prepaid postcard.

  
Kirk E. Pastorian  
Registration No. 48,756  
Agent of Record

R:\DOCS\KEP\KEP-1237.DOC  
103101

# BUNDESREPUBLIK DEUTSCHLAND



RECEIVED

JAN 14 2002

Technology Center 2600

## Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

**Aktenzeichen:** 198 31 758.1

**Anmeldetag:** 15. Juli 1998

**Anmelder/Inhaber:** Jerini BIO TOOLS GmbH, Berlin/DE

**Bezeichnung:** Ligandenbestimmung für Proteine

**IPC:** G 01 N, C 12 Q

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 17. Oktober 2001  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag

HoiB

Jerini BIO TOOLS GmbH, 12489 Berlin

---

## Ligandenbestimmung für Proteine

---

Die vorliegende Erfindung betrifft ein Verfahren zur Bestimmung von Liganden für Proteine gemäß den Merkmalen des Patentanspruches 1.

In der Biochemie versteht man unter Liganden biologische aktive Substanzen meist niedermolekularer Art, die durch Bindung an eine spezifische Bindungsstelle eines Makromoleküls eine bestimmte Wirkung auf das Makromolekül ausüben. Bei den hier betreffenden Makromolekülen kann es sich um Enzyme, Rezeptoren, DNA, RNA usw. handeln.

Durch Bindung des Liganden an das Makromolekül können beispielsweise der katalytische Umsatz eines Enzyms, die Aktivierung bzw. Inaktivierung eines Enzyms sowie Konformationsänderungen von Makromolekülen bewirkt werden.

In der pharmazeutischen Industrie werden bisher zwei Strategien zur Identifizierung biologischer aktiver Substanzen, d.h. Liganden, angewandt.

Die Unternehmen verfügen in der Regel über große Substanzsammlungen vieler verschiedener Einzelverbindungen. Diese Substanzen werden in biologischen Systemen, z.B. Zellassays, mittels Hochdurchsatzverfahren in Form von Pipettierstraßen mit automatischer Auswertung auf bestimmte Aktivitäten gete-

stet. Treffer bei diesen Verfahren sind allerdings nur zufällig, sie treten aber mit bestimmten Wahrscheinlichkeiten auf.

Die Alternative dazu ist eine andere Strategie, die mit Hilfe von Computern durchgeführt wird. Unter Berechnung der Kräfte zwischen Molekülen werden am Computer Verbindungen, die an bestimmte Proteinoberflächen binden sollen, virtuell erstellt und dann erst synthetisiert. So werden im Gegensatz zum obigen Verfahren weniger Substanzen synthetisiert und getestet. Es werden auch virtuelle Substanzbibliotheken von Molekülen, die nicht als Substanz vorliegen müssen, im Dockingverfahren am Computer auf eine Bindung an eine bestimmte Proteinoberfläche getestet. Wiederum werden dann nur die Treffer synthetisiert und in biologischen Testsystemen eingesetzt. Verfahren dieser Art sind bereits in den US-Patenten 5,495,423, 5,579,250 und 5,612,895 beschrieben worden.

In der Praxis wurden auch Kombinationen der oben beschriebenen Verfahren angewendet.

Bei diesen Verfahren wurden aber keine in der Natur vorkommenden Interaktionen ausgenutzt. Des weiteren sind viele bekannte Verfahren der Zufälligkeit ausgesetzt und müssen sich oftmals auf die virtuelle Beobachtung stützen. Dieses führt zu beträchtlichen Zeitverlusten und Ungenauigkeiten.

Es ist daher Aufgabe der vorliegenden Erfindung, ein Verfahren zur Verfügung zu stellen, bei dem die Liganden für Proteine schnell und zuverlässig bestimmt werden können.

Die Aufgabe wird mit einem Verfahren gemäß Patentanspruch 1 gelöst.

Die Unteransprüche betreffen bevorzugte Ausführungsformen des erfindungsgemäßen Verfahrens.

Die Erfindung betrifft weiterhin Liganden, die nach dem erfindungsgemäßen Verfahren hergestellt sind.

Das erfindungsgemäße Verfahren zur Bestimmung von Liganden für Proteine umfaßt folgende Schritte:

- a) Bestimmung der Sekundärstrukturelemente eines gegebenen Proteins, die den Bindungsort für den Liganden bilden;
- b) Zerlegung der molekularen Oberfläche eines gegebenen Proteins in molekulare Oberflächenteile;
- c) Bestimmen von ähnlichen Flächen zu denjenigen Elementen, die die zu bestimmende Bindungsregion für den Liganden definieren, wobei die gefundenen molekularen Oberflächenteile ein komplementäres Nachbarelement besitzen;
- d) Koordinatentransformation des gefundenen molekularen Oberflächenteils mit Nachbarelement auf ein Ausgangselement bei einem rms-Wert unterhalb von 2Å und
- e) Beurteilung der Paßfähigkeit des Liganden gemäß lokaler Packungsdichte.

Der Ablauf des erfindungsgemäßen Verfahrens wird anhand des in Fig. 1 gezeigten Fließdiagramms erläutert.

Vorzugsweise wird das erfindungsgemäße Verfahren auf der Grundlage einer Datenbank durchgeführt. Es hat sich als zweckmäßig erwiesen, die Datenbank "Dictionary of Interfaces in Proteins (DIP)" zu verwenden, die in J.Mol.Biol. (noch nicht veröffentlicht) beschrieben ist. Die Datenbank DIP stellt Flächen zwischen Sekundärstrukturen (SSE) aller strukturell bekannten Proteine zur Verfügung. Diese Interfaces bestehen aus zwei Atommengen (Patches), die Teile von benachbarten Sekundärstrukturen sind und zusammen den Kontakt dieser beiden Strukturen ausmachen.

Bei der Bestimmung von Liganden, dem sogenannten "drug design", stellt sich die Frage, welche chemische Verbindung zu einer gegebenen Proteinstruktur paßt. Erfindungsgemäß wer-

den die Sekundärstrukturelemente eines gegebenen Proteins, die den Bindungsort für den Liganden bilden, bestimmt. Dann wird die molekulare Oberfläche eines gegebenen Proteins zunächst in molekulare Oberflächenteile (MSP = molecular surface patches) zerlegt. Zu denjenigen Elementen, die potentiell die Bindungsregion definieren, werden beispielsweise aus der oben beschriebenen Datenbank ähnliche Flächen herausgesucht. Als Nebenbedingung wird bei dem Screening auf Ähnlichkeit verlangt, daß die gefundenen MSPs bereits ein komplementäres Nachbarelement besitzen. Eine Transformation, beispielsweise Koordinatentransformation, des gefundenen MSP mit Nachbarelement auf das Ausgangselement ist dann aussichtsreich, wenn der rms-Wert (mittlerer Fehler) unterhalb von 2Å liegt. Vorzugsweise liegt der Wert bei 1,5 Å. Für die Beurteilung der Paßfähigkeit des Liganden gegenüber dem Original hat sich die lokale Packungsdichte nach Goede et al. bewährt.

Bei dem erfindungsgemäßen Verfahren sollen die Außenflächen der Sekundärstrukturen bestimmt werden. Die Außenflächen, die den Kontakt herstellen, sind die molekularen Oberflächenteile (MSP). Ähnliche molekulare Oberflächenteile werden überlagert. Nach der Koordinatentransformation liegen die gefundenen molekularen Oberflächenteile auf Atomen des Bindungsortes. Die besten potentiellen Liganden bilden die Leitverbindung. Der Vergleich der besten potentiellen Liganden mit einem bekannten Ausgangsprotein plus Ligand erfolgt als letztes.

Somit erfolgt erfindungsgemäß die Bestimmung eines komplementären Bindungspartners dadurch, indem ähnliche Elemente bestimmt werden, die bereits einen Bindungspartner besitzen.

Wenn die Liganden, wobei es sich um Sekundärstrukturelemente aus ca. 10 Aminosäuren handelt, bestimmt sind, müssen sie für den Einsatz als Pharmakon noch optimiert werden, da Peptide aus den natürlichen L-Aminosäuren vielen Anforderungen nicht entsprechen.

Es gibt experimentelle Verfahren zur synthetischen Umwandlung von Peptiden in Peptidmimetika, z.B. Peptoide, die vom pharmakologischen Standpunkt aus betrachtet, häufig wesentlich günstigere Eigenschaften aufweisen. Die Verbindungen durchlaufen dabei in der Regel verschiedene Optimierungszyklen, in denen die Moleküle auch wirklich als Substanzen vorliegen.

Eine weitere Möglichkeit, Leitverbindungen zu finden, besteht in der Suche in Datenbanken niedermolekularer Verbindungen. In diesem Falle werden die Koordinaten des entsprechenden passfähigen Peptids oder Teile davon verwendet, um in einer entsprechenden Datenbank nach dem angegebenen Überlagerungsverfahren (Vergleichsverfahren) zu suchen. Damit ist es möglich, völlig unabhängig von der peptoiden Grundstruktur Leitverbindungen zu finden.

Das erfindungsgemäße Verfahren zur Bestimmung von Liganden wird vorzugsweise für die aktiven Zentren von Enzymen beschrieben. Das Verfahren ist allerdings auch auf andere Makromoleküle (Proteine, DNA, RNA) übertragbar, sofern sie geeignete Oberflächen besitzen. Folgende Anwendungsgebiete kommen beispielsweise in Betracht:

- \* Bindungs- und/oder Nachweismoleküle in diagnostischen Assays
- \* Lebensmittelindustrie: Suche von Liganden für Geschmacksrezeptoren und Verwendung als Geschmackszusatzstoff
- \* Biotechnologie: Moleküle für die Affinitätsreinigung
- \* Proteine, die im therapeutischen Bereich gebunden werden müssen:
  - Enzyme, Rezeptoren, DNA, RNA
  - Zytokine oder Wachstumsfaktoren und ihre Rezeptoren, insbesondere diejenigen, die der Stoffwechselregulation dienen
  - Zelladhäsionsproteine und ihre Rezeptoren
  - Proteine der Signaltransduktionswege und ihre Bindungspartner
  - zytosolische Rezeptoren, Steroidrezeptoren

Proteine der Blutgerinnung  
 Neurotransmitter und ihre Rezeptoren  
 Proteine der Stoffwechselwege  
 Proteine der Replikation, Transkription und  
 Translation  
 Proteine von Krankheitserregern (Bakterien,  
 Viren, eukaryotische Einzeller, Parasiten)

Das erfindungsgemäße Verfahren läßt sich auch zur Bestimmung von Proteinstrukturen anwenden. Es ist nicht auf alleinige Sequenzähnlichkeit angewiesen, sondern verwendet Strukturähnlichkeit von molekularen Grenzflächen von Sekundärstrukturelementen zur Vorhersage ihrer Wechselwirkungspartner. Dabei wird der Tatsache Rechnung getragen, daß gleiche (ähnliche) Grenzflächen auch bei unterschiedlichen Sequenzen entstehen können.

Beispielhaft soll im folgenden das Verfahren der Proteinstrukturbestimmung in seinen Schritten geschildert werden.

Im ersten Schritt wird eine gegebene Primärstruktur in ihrer vollen Länge in eine repetitive Sekundärstruktur

"gewickelt". Dies bedeutet, daß mit Standard  $\phi$ ,  $\varphi$  und  $\chi$ -Winkeln  $\beta$ -Faltblätter bzw.  $\alpha$ -Helices über die volle Länge der Primärstruktur errechnet werden.

Im zweiten Schritt werden die entstandenen molekularen Grenzflächen dieser Sekundärstrukturelemente geclustert und mit einem artifiziellen neuronalen Netz bewertet, dessen Eingangsdaten sich aus den molekularen Oberflächen der geclusterten Strukturelemente ergeben. Die Bewertung erfolgt mit dem Ziel, einerseits zu bestätigen, ob überhaupt mit der gegebenen Primärstruktur molekulare Oberflächen bei dem Sekundärstrukturelement zu gestalten sind, die repräsentativ für das gegebene Strukturelement sind. Wenn nicht, wird die Sekundärstruktur verworfen. Damit ergibt sich ein neues Verfahren der Sekundärstrukturvorhersage. Das neuronale Netz wird anhand der bekannten Proteinstrukturen trainiert.



Alternativ zur allgemeinen Strukturbildung auf der Grundlage von Standard  $\phi$ ,  $\psi$  und  $\chi$ -Winkeln für Helices bzw. Faltblättern können bekannte Sekundärstrukturvorhersagealgorithmen angewendet werden, so daß das vorgenannte Verfahren nur auf die vorhergesagten Strukturen (Teile der Sequenz) angewandt wird. In einem weiteren Schritt werden die gefundenen Cluster, die Kontakt zu einem bestimmten Sekundärstrukturelement (oder Solvent) haben, verwendet, um in der Datenbank DIP gleiche oder ähnliche molekulare Oberflächen und deren Nachbarn zu suchen. Dies geschieht mit dem weiter oben schon beschriebenen bias-freien Überlagerungsalgorithmus für atomare Sets.

Aus dem vorgenannten Arbeitsschritt ergeben sich eine Reihe von molekularen Oberflächen (MPS), deren Partnerelement sicher bzw. weniger sicher (Variantenplanung) festliegt. Wird dabei "nicht-Solvent" vorhergesagt, versucht ein einfacher Docking-Algorithmus im dritten Schritt eine passende Fläche in anderen Sekundärstrukturelementen als dem direkt betrachteten zu lokalisieren. Der einfache Docking-Algorithmus beruht auf der Tatsache, daß molekulare Grenzflächenpartner zwischen Sekundärstrukturen innerhalb eines gegebenen Abstandes der beiden Schwerpunkte bzw. einem bestimmten Winkel der ausgezeichneten Richtung gesucht werden kann. Die Qualität der Paßfähigkeit wird mit Hilfe der molekularen Dichtebestimmung überprüft (Goede et al. Journal of Computational Chemistry, Bd. 18, Nr. 9, S. 1114ff, 1997). Liegen die potentiellen Partner fest, wird in einem vierten Schritt die prinzipielle Faltbarkeit unter Einhaltung aller vorhergesagten Nachbarschaften (Solvent, Helix-Helix, Helix-coil, Helix-extended) überprüft und die allgemeine Faltung oder mehrere Varianten von der gegebenen Sequenz angenommen.

Das folgende Beispiel soll das erfindungsgemäße Verfahren erläutern.

## Beispiel

### Inhibitor-Design für das Proteasom

Ausgehend von einem Bindungsort einer aktiven Untereinheit des Proteasoms bei Hefe werden die Sekundärelemente bestimmt, die den Bindungsort bilden. Es stellt sich heraus, daß fünf Elemente beteiligt sind, wobei zwei größere Elemente den Bindungsort bestimmen. Anschließend werden die Außenflächen dieser Sekundärstrukturen bestimmt. Mit den Teilen der Außenflächen, die den Kontakt ausmachen und 12 bis 22 Atome umfassen, werden in der DIP-Datenbank ähnliche MSPs gesucht. Die ähnlichen MSPs von einer bestimmten Mindestgüte, wobei mindestens 70% der Atome überlagert sind und der rms-Wert  $1,0\text{\AA}$  beträgt, werden mit den Ausgangsflächen überlagert, wobei die Aminosäuren, die die Gegenseite der MSPs bilden bei der Koordinatentransformation der MSPs einbezogen werden. Nach der Koordinatentransformation liegen die gefundenen MSPs auf den Atomen des Bindungsortes, die Gegenseiten der MSPs in der Bindungstasche.

Die Gegenseiten der gefundenen MSPs, die die potentiellen Liganden darstellen, werden dahingehend überprüft, ob sie die Bindungstasche ausfüllen und ob die Abstände zu den Atomen der Bindungstasche groß genug sind. Dafür wird die lokale Dichte in der Bindungstasche berechnet. Die besten potentiellen Liganden bilden die Leitverbindungen.

Ein Vergleich der zehn besten potentiellen Liganden mit einer Proteasom-Struktur des Archebakteriums, die mit Ligand vorliegt, ergibt, daß die Hauptkette von einer der auf diese Weise berechneten Strukturen völlig identisch zu dem bekannten Inhibitor des Proteasoms des Archebakteriums ist.

## PATENTANSPRÜCHE

1. Verfahren zur Bestimmung von Liganden für Proteine, das folgende Schritte umfaßt:

- a) Bestimmung der Sekundärstrukturelemente eines gegebenen Proteins, die den Bindungsort für den Liganden bilden;
- b) Zerlegung der molekularen Oberfläche des Proteins in molekulare Oberflächenteile;
- c) Bestimmen von ähnlichen Flächen zu denjenigen Elementen, die die zu bestimmende Bindungsregion für den Liganden definieren, wobei die gefundenen molekularen Oberflächenteile ein komplementäres Nachbarelement besitzen;
- d) Koordinatentransformation des gefundenen molekularen Oberflächenteils mit Nachbarelement auf ein Ausgangselement bei einem rms-Wert unterhalb von 2Å und
- e) Beurteilung der Paßfähigkeit des Liganden gemäß lokaler Packungsdichte.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Außenflächen der Sekundärstrukturen bestimmt werden.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß die Außenflächen, die den Kontakt herstellen, die molekularen Oberflächenteile sind.

4. Verfahren nach einem der vorangegangenen Ansprüche, dadurch gekennzeichnet, daß die ähnlichen molekularen Oberflächenteile mit den Ausgangsflächen überlagert werden.

5. Verfahren nach einem der vorangegangenen Ansprüche, dadurch gekennzeichnet, daß nach der Koordinatentransformation die gefundenen molekularen Oberflächenteile auf Atomen des Bindungsortes liegen.

6. Verfahren nach einem der vorangegangenen Ansprüche, dadurch gekennzeichnet, daß die besten potentiellen Liganden die Leitverbindung bilden.

7. Verfahren nach einem der vorangegangenen Ansprüche, dadurch gekennzeichnet, daß die besten potentiellen Liganden mit einem bekannten Ausgangsprotein plus Ligand verglichen werden.
8. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Liganden in Form von Peptiden bestimmt werden.
9. Verfahren nach Anspruch 8, dadurch gekennzeichnet, daß das Peptid etwa zehn Aminosäuren umfaßt.
10. Verfahren nach Anspruch 9, dadurch gekennzeichnet, daß das Peptid anschließend in ein Peptidmimetikum umgewandelt wird.
11. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Proteine Enzyme sind.
12. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß der rms-Wert 1,5Å beträgt.
13. Verfahren nach einem der vorangegangenen Ansprüche, dadurch gekennzeichnet, daß es zur Strukturbestimmung von Proteinen angewendet wird.
14. Verwendung eines nach den Ansprüchen 1 bis 12 hergestellten Ligands zur Herstellung eines Pharmakons.

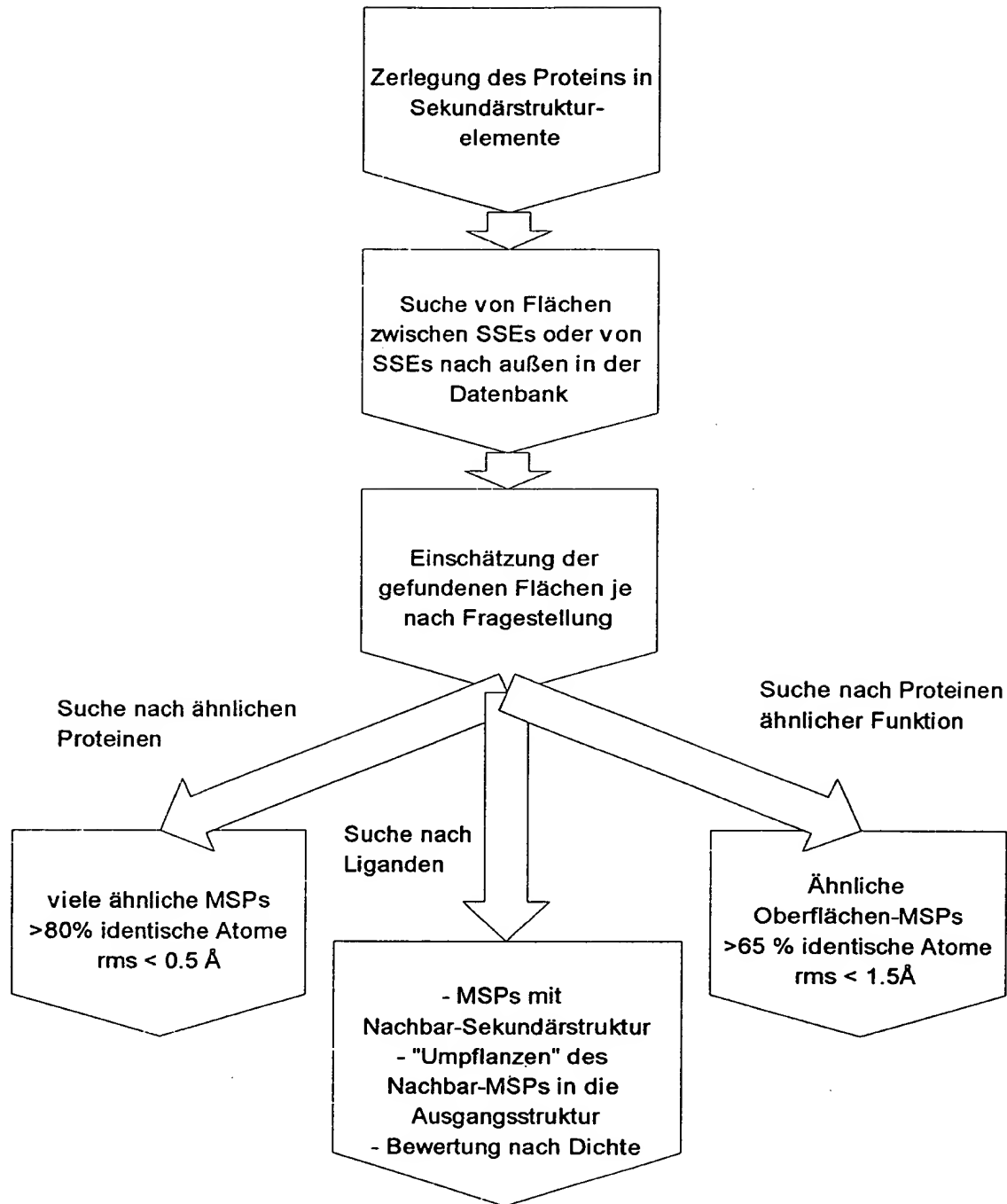


Fig. 1



Creation date: 01-06-2004  
Indexing Officer: LTOUCH - LY TOUCH  
Team: OIPEBackFileIndexing  
Dossier: 09772538

Legal Date: 03-29-2002

No.	Doccode	Number of pages
1	IMIS	1

Total number of pages: 1

Remarks:

Order of re-scan issued on .....